

摘要

本文目的是將實體機電系統簡化後導入虛擬環境並證明簡化後訓練的強化學習能算法能在模擬環境中應用。

將實體冰球機的機電系統的雙自由度簡化成一個自由度導入 CoppeliaSim 模擬環境透過 Remote API 控制環境中的冰球機移動，OpenCV 來處理影像提供強化學習訓練和訓練後實際控制的輸入，強化學習訓練利用 OpenAI Gym 的 Pong Game 測試適合的訓練參數，再將此算法套用到 CoppeliaSim 的場景中進行訓練。

該研究為了證實相同的訓練算法能套用到不同真實程度的環境中進行訓練（，相同訓練參數套用到模擬環境中進行訓練，依舊能訓練出像樣的對打系統。）

關鍵字: 類神經網路、強化學習、CoppeliaSim、

OpenAI Gym

Chapter 1

介紹

近年來硬體技術、軟體、自動求導等技術快速發展起來，再次帶起機器學習的發展，促使機器學習與各領域結合的應用越來越廣泛，在機電系統採用強化學習是為了讓機電系統的控制達到最佳化。本研究利用強化學習優化冰球機的對打系統，並測試相同演算法運用在 2D 與 3D 模擬環境可行性。

本研究分成兩大部分，第一部分簡化冰球機並運用 OpenAI Gym 的 Atari Pong-v0 測試較合適的訓練參數，在 2D 環境中進行強化學習的對打訓練。第二部分將實體系統簡化後導入 CoppeliaSim 模擬環境導入測試完成的訓練參數進行虛擬訓練，進行算法導入到實體機前的測試。

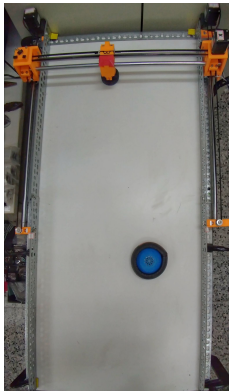


Figure 1.1: 實體的冰球機

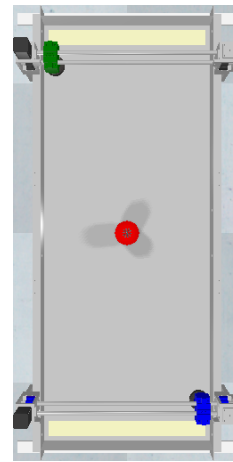


Figure 1.2: 虛擬環境簡化後的冰球機

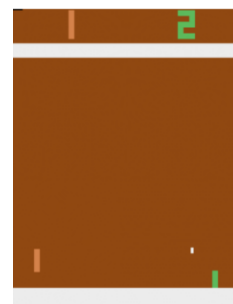


Figure 1.3: Gym 的 Pong game

Chapter 2

Literature Review

Atari Video Computer System 是 Atari Inc 在 1977 年 9 月 11 日發行的一款家用遊戲機，直到 1982 年更名為 Atari 2600，Pong 是第一個主要發行的遊戲。OpenAI 是一個人工智慧研究實驗室，研究著重於強化學習，Gym 具有多種不同的環境，並以標準化定義環境，其中包括 Atari 2600 的遊戲、Classic control、Box2D 等環境供機器學習訓練使用。

冰球機與 Pong 都是透過玩家的眼睛觀察球的移動並透過操控擊錘進行攻擊與防守，並依照對手的球路來採取不同的策略，兩者最大的差異是擊錘的移動自由度，冰球機的擊錘自由度為 2，則 Pong 的自由度為 1，

簡化成單一自由度並透過 Pong 來訓練

Pong 是一個 2D 的運動遊戲，模擬乒乓球對打，藉由移動擊錘進行對打。強化學習的訓練是透過影像處理，將擊錘及球的位置過濾出來，比較兩幀間擊錘與球的移動，透過強化學習判斷擊錘的移動，訓練一個。

Chapter 3

設計流程

3.1 研究動機

機器學習與各領域結合的應用越來越廣泛，在機電系統採用強化學習是為了讓機電系統的控制達到最佳化。本專題以實體的冰球機(圖.1.1)之機電系統作為訓練模型，將實體機器轉移到虛擬環境(圖.1.2)進行模擬，為了找到適合的演算方式，因此將模型簡化(圖.1.3)後再進行測試各種算法的優劣，透過不斷的訓練來得到一個優化過的對打系統。

3.2 研究目的與方法

本研究分三大部分，第一運用 OpenAI Gym 裡內建編譯的 ATARI 2600 遊戲 Pong-v0，來作為訓練環境，加上強化學習的理論，測試不同演算法以訓練出最佳化的對打系統，第二換為 CoppilaSim 模擬環境並套用訓練程式，成為優化的對打機電系統。第三則是透過架設伺服器與虛擬環境結合。

利用 Gym 的訓練環境來測試不同的算法所得到的訓練結果，比較不同算法、參數間的差異，並找出較適合 Pong game 的算法、參數，循序漸進提高環境的真實程度，來減少一開始就是以實體的方式測試所帶來硬體、程式、時間和金錢等成本。

將 Gym 的訓練環境轉換到 CoppilaSim 模擬環境，利用貼近真實的模擬環境來修正在純程式的架構(Gym)與真實環境間的誤差，雖然 CoppilaSim 模擬環境與真實環境仍有些微的落差，兩者相比 CoppilaSim 的環境已非常貼近真實了，拉近了虛擬與現實間的距離，提高了實用性的價值。這部分還有利利用 OpenCV 進行影像處理並撰寫輔助對打程式來協助玩家預判球的移動路徑或彈射位置。影像處理除了應用在輔助對打上，最主要是應用在訓練強化學習所需的輸入訓練影像。

再透過架設伺服器與虛擬環境結合：讓虛擬環境的影像透過網伺服器串流影像供使用者遠端進行操控虛擬環境的擊錘進行打球，或是提供多人進行觀看對打影像。

3.3 未來展望

此專題希望能利用現有完成的機械學習的算法，能發展成虛擬訓練，再將訓練完的機器學習應用到虛擬環境或是實體機電系統，並透過伺服器將影像串流提供玩家網頁介面進行遠端操控，同時提供多人觀看及時的比賽影像，將整個冰球機的控制和使用者間有更完善串聯，機電系統的部分達到最優化控制和虛實整合的應用。